REMARKS

Upon entry of the foregoing amendments, claims 1–4, 10–14, 16, 18, 19, 25 and 26 are pending. Applicants have amended claims 3 and 4 to correct antecedent basis. Thus, claims 3 and 4 now recite "polynucleotides." Basis for these amendments can be found in the specification as originally filed, and in particular, in originally filed claims 1 and 2. Applicants have amended claim 13 to correct improper dependency. Thus, claim 13 now recites "the host cell of claim 12." Applicants have amended claim 19 to correct clerical errors and improper dependency. Thus, claim 19 now recites "the vaccine vector of claim 16." Applicants have amended claim 25 to correct clerical errors. Applicants have added new claims 38 and 39 to more accurately claim the present invention. New claims 38 and 39 further limit the type of host cells to be used in the present invention. Specifically, new claims 38 and 39 are directed to mammalian cells and human cells respectively. Basis for these amendments can be found in the specification as originally filed, and in particular, at pg. 14, lines 9–14; pg. 15, line 31; and in originally filed claim 17.

These amendments add no new matter.

Co–Pending Applications

Applicants acknowledge the Examiner's request regarding copies of claims, correlated with the serial number of the case in which they appear, for each pending application directed to the pending subject matter.

Applicants respectfully submit that acquiescing to the Examiner's request of providing copies of pending claims would be unduly onerous and expensive for Applicants. Applicants have **no obligation to provide the Examiner with copies of claims in any pending applications**. In fact, Applicants submit that, under M.P.E.P. §804, it is the Examiner's burden to identify any potential statutory and/or "obvious-type" double patenting rejections. *See*, M.P.E.P. §804(I)(B), which states:

> Occasionally, the **examiner** becomes aware of two copending applications filed
> by the same inventive entity, or by different inventive entities having a common

inventor, and/or that are filed by a common assignee that would raise an issue of
double patenting if one of the applications became a patent.

However, Applicants agree to provide the Examiner copies of the requested claims upon
determination of allowable subject matter in the present application.

Drawings

Applicants acknowledge the Draftsperson's objection to the Drawings filed in the present
application.

Applicants will provide formal drawings upon determination of allowable subject matter
in the present application.

Oath/Declaration

Applicants acknowledge the Examiner's objection to the Declaration as filed in the
present application. Applicants have enclosed herewith an executed Declaration. Accordingly,
Applicants believe that the present objection is now moot.

Specification

The Examiner has objected to the Specification as having various informalities.

Applicants have amended the Specification to correct these informalities. Accordingly,
Applicants believe that the present objection is now moot.

Provisional Double Patenting Rejection over Co-pending Applications

The Examiner has provisionally rejected claims 1–4, 10–14, 16–19, 25 and 26 under 35
U.S.C. §101 as claiming the same invention as that of claims 1, 20, 21, 25–28, 32–35, 41 and 42
of co-pending application U.S.S.N. 09/376,770 ("'770").

Applicants traverse. 37 CFR §1.78(b) provides that when two or more applications filed
by the same Applicant contain conflicting claims, elimination of such claims from all but one
application may be required in the absence of good and sufficient reason for their retention

during pendency in more than one application. However, the M.P.E.P. clearly states that "[t]his paragraph is appropriate only when the conflicting claims are **patentably distinct**. See MPEP §822. *See also*, In re Zickendraht, 319 F.2d 225, 138 USPQ 22 (CCPA 1963) (where the Court held that the doctrine is well established in that claims in different applications need be more than merely different in form or content and that **patentable distinction must exist to entitle applicants to a second patent**).

The '770 application is a co-pending application having the both the same inventive entity and assignee. However, both applications are directed to distinct sequences encoding *different Chlamydia* polypeptides. Applicants have attached herewith an alignment comparison between the polypeptide of the present application with those disclosed by U.S.S.N. 09/376,770 (Exhibits A and B). The comparisons clearly show that the polypeptides according to the present invention are **completely different** from those disclosed in the '770 application.

Accordingly, in light of the arguments above, Applicants respectfully request reconsideration and withdrawal of the present provisional double patenting rejection.

35 U.S.C. §112, First Paragraph Rejections

The Examiner has rejected claims 1–4, 10–14, 16–19, 25 and 26 under 35 U.S.C. §112, first paragraph as containing subject matter which was not described in the Specification in such a way as to enable one skilled in the art to make and/or use the invention, stating that the Specification is not enabled for a polynucleotide encoding a polypeptide having a sequence that is "at least 75% homologous" to SEQ ID NO:2 and "functional fragments thereof." Specifically, the Examiner has alleged that:

> Without a clear and unambiguous description of how to perform the comparison, the scope of the claims can not be envisaged (*sic*). Without a specific disclosure of the parametric values used in the algorithm, the sequence identity between two sequences has no common meaning within the art and therefore, one of ordinary skill in the art cannot be sure of the sequences embraced by the claims and would not be able to make and use those polynucleotide or polypeptide sequence homologs as recited in the instant claims, without undue experimentation.

See e.g., December 13, 2000 Office action at pg. 5.

 Applicants traverse. Methods and computational programs for conducting sequence comparisons for homology are well known within the art. Homology is typically measured using **sequence analysis software** (*e.g.*, Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705). Similar amino acid sequences are aligned to obtain the maximum degree of homology (*i.e.*, identity). To this end, it may be necessary to artificially introduce gaps into the sequence. Once the optimal alignment has been set up, the degree of homology (*i.e.*, identity) is established by recording all of the positions in which the amino acids of both sequences are identical, relative to the total number of positions.

 The Specification also describes a preferred method for determining amino acid similarities, including amino acid sequence homology:

> One particularly preferred method of determining amino acid similarities is the PAM250 matrix described in Dayhoff et *al.,* 5 ATLAS OF PROTEIN SEQUENCE AND STRUCTURE 345-352 (1978 & Supp.), incorporated by reference herein. A similarity score is first calculated as the sum of the aligned pairwise amino acid similarity scores. Insertions and deletions are ignored for the purposes of percent homology and identity. Accordingly, gap penalties are not used in this calculation. The raw score is then normalized by dividing it by the geometric mean of the scores of the candidate compound and the reference sequence. The geometric mean is the square root of the product of these scores. The normalized raw score is the percent homology.

See e.g., Specification at pg. 9, lines 15–23. Furthermore, the Specification specifically and clearly defines a "homologous amino acid sequence" as:

> an amino acid sequence that differs from an amino acid sequence shown in SEQ ID NO: 2, only by one or more conservative amino acid substitutions, or by one or more non-conservative amino acid substitutions, deletions, or additions located at positions at which they do not destroy the specific antigenicity of the polypeptide. Preferably, such a sequence is at least 75%, more preferably 80%, and most preferably 90% identical to an amino acid sequence shown in SEQ ID NO: 2. Homologous amino acid sequences include sequences that are identical or substantially identical to an amino acid sequence as shown in SEQ ID NO: 2.

Id at pg. 8, lines 21–29. The Specification also describes amino acid sequence identity and conservative substitutions as:

> By "amino acid sequence substantially identical" is meant a sequence that is at least 90%, preferably 95%, more preferably 97%, and most preferably 99% identical to an amino acid sequence of reference and that preferably differs from the sequence of reference, if at all, by a majority of conservative amino acid substitutions. Conservative amino acid substitutions typically include substitutions among amino acids of the same class. These classes include, for example, (a) amino acids having uncharged polar side chains, such as asparagine, glutamine, serine, threonine, and tyrosine; (b) amino acids having basic side chains, such as lysine, arginine, and histidine; (c) amino acids having acidic side chains, such as aspartic acid and glutamic acid; and (d) amino acids having nonpolar side chains, such as glycine, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan, and cysteine.

Id at pg. 9, lines 8–14. Thus, Applicants assert that one of ordinary skill in the art could make and use such homologous sequences in, for example, vaccines and diagnostic reagents. Accordingly, Applicants respectfully request reconsideration and withdrawal of the present rejection.

35 U.S.C. §112, Second Paragraph Rejections

The Examiner has rejected claims 3, 4, 13, 17 and 19 under 35 U.S.C. §112, second paragraph, as being indefinite for improper antecedent basis.

Applicants have amended claims 3 and 4 to correct antecedent basis. Thus, claims 3 and 4 now recite "polynucleotides." Basis for these amendments can be found in the specification as originally filed, and in particular, in originally filed claims 1 and 2. Applicants have amended claim 13 to correct improper dependency. Thus, claim 13 now recites "the host cell of claim 12." Applicants have amended claim 19 to correct improper dependency. Thus, claim 19 now recites "the vaccine vector of claim 16." Applicants have added new claims 38 and 39 to more accurately claim the present invention. New claims 38 and 39 further limit the type of host cells to be used in the present invention. Specifically, new claims 38 and 39 are directed to mammalian cells and human cells respectively. Basis for these amendments can be found in the specification as originally filed, and in particular, at pg. 14, lines 9–14; pg. 15, line 31; and in originally filed claim 17.

In light of the above amendments, Applicants believe that the present rejections are now moot. Accordingly, Applicants respectfully request reconsideration and withdrawal of the present rejections.

35 U.S.C. §102 Rejection

The Examiner has rejected claims 1–4, 10–13, 16–19, 25 and 26 under 35 U.S.C. §102(b) as being anticipated by Longbottom *et al.* (**Longbottom**). Specifically, the Examiner has stated that **Longbottom** teaches "*chlamydial* genes or sequences coding for highly immunogenic protein fragments comprising 8 or 9 amino acid residues."

Applicants traverse. **Anticipation requires the disclosure in a single prior art reference of each element of the claim under consideration.** *W.L. Gore & Associates v. Garlock, Inc.,* 220 USPQ 303, 313 (Fed. Cir. 1983), *cert. denied,* 469 U.S. 851 (1984); *Connell v. Sears Roebuck & Co.,* 220 USPQ 193, 198 (Fed. Cir. 1983); *Verdegaal Bros. v. Union Oil Co. of California,* 2 USPQ2d 1051, 1053 (Fed. Cir. 1987); *In re Spada,* 15 USPQ2d 1655 (Fed. Cir. 1990); MPEP § 2131. **"There must be no difference between the claimed invention and the reference disclosure, as viewed by a person of ordinary skill in the field of the invention."** *Scripps Clinic & Research Foundation v. Genentech Inc.,* 18 USPQ 2d 1001, 1010 (Fed. Cir. 1991).

Although both applications are directed to *Chlamydia* polypeptides, **Longbottom** discloses distinct sequences encoding *different Chlamydia* polypeptides. Applicants have attached herewith an alignment comparison between the polypeptides of the present application with those disclosed by **Longbottom** (Exhibit C1 to C4). Applicants have highlighted the regions that the Examiner has suggested would be identical between the present application and **Longbottom** . As can be seen by the highlighted regions, the rejection seems to be based upon that of identical fragments within the polynucleotides and/or polypeptides of the present invention, **not on the *entire* polynucleotide or polypeptide.** Furthermore, the comparisons clearly show that the polynucleotides and polypeptides according to the present invention are, in

fact, **different** from those disclosed in **Longbottom**. Thus, Applicants assert that the present application is not anticipated by **Longbottom.**

Accordingly, in light of the arguments above, Applicants respectfully request reconsideration and withdrawal of the present rejection.

CONCLUSION

On the basis of the foregoing amendments and remarks, Applicants respectfully submit that the pending claims are in condition for allowance. If there are any questions regarding these amendments and remarks, the Examiner is encouraged to contact the undersigned at the telephone number provided below.

Respectfully submitted,

Ivor R. Elrifi, Reg. No. 39,529
Michel Morency, Reg. No. P-50,183
Attorneys for Applicants
c/o MINTZ, LEVIN
One Financial Center
Boston, Massachusetts 02111
Tel: (617) 542-6000
Fax: (617) 542-2241

Dated: December 4, 2001

VERSION WITH MARKINGS TO SHOW CHANGES MADE

In the Specification:

The paragraph beginning at line 6, pg. 8 was amended as follows:

-- [FIG. 1] <u>FIG. 1A to 1I</u> shows the nucleotide sequence (top sequence) and the deduced amino acid sequence (bottom sequence) of the full length *98 kDa putative outer membrane protein* gene (SEQ ID NO: 1) and the processed sequence from *Chlamydia pneumoniae* (SEQ ID NO: 2). --

The paragraph beginning at line 9, pg. 8 was amended as follows:

-- [FIG. 2] <u>FIG. 2A to 2H</u> shows the restriction enzyme analysis of nucleotide sequence encoding the *C. pneumoniae 98 kDa putative outer membrane protein* gene. --

The paragraph beginning at line 13, pg. 16 was amended as follows:

A recombinant expression system can be selected from prokaryotic and eukaryotic hosts. Eukaryotic hosts include yeast cells (*e.g., Saccharomyces cerevisiae* or *Pichia pastoris*), mammalian cells (*e.g.,* COS1, NIH3T3, or JEG3 cells), arthropods cells (*e.g., Spodoptera frugiperda* (SF9) cells), and plant cells. Preferably, a prokaryotic host such as *E. coli* is used. Bacterial and eukaryotic cells are available from a number of different sources to those skilled in the art, *e.g.,* the American Type Culture Collection [(ATCC; Rockville, Maryland)] <u>(ATCC; 10801 University Boulevard, Manassas, VA 20110-2209)</u>.

In the Claims:

3.   (Amended) The <u>poly</u>nucleotide of claim 2 wherein the fusion polypeptide is a heterologous signal peptide.

4.   (Amended) The polynucleotide of claim 2 wherein the polynucleotide encodes a functional fragment of the polypeptide having the SEQ ID NO: 2.

13.   (Amended) The host cell of claim 12[0], wherein said host cell is a prokaryotic cell.

19.   (Amended) A pharmaceutical composition, comprising an immunologically effective amount of the vaccine vector of claim 16[4].

25.   (Amended) A polynucleotide probe reagent capable of detecting the presence of *Chlamydia* in a biological material, comprising a polynucleotide that hybridizes to the polynucleotide of claim 1 under stringent conditions.

--38.   (New) The host cell of claim 14, wherein said eukaryotic cell is a mammalian cell.

39.   (New) The host cell of claim 38, wherein said mammalian cell is a human cell.--

TRA 1460374v2

EXHIBIT C - The Alignment of LONGBOTTOM to CPN 100396 (SEQ ID NO:1)

Alignment 1

```
GAP of: cpu72499.seq  check: 4475  from: 1  to: 4926

WPDEF       Chlamydophila abortus putative outer membrane protein gene, partial
LOCUS       CPU72499     4926 bp    DNA              BCT        13-MAY-1999
DEFINITION  Chlamydophila abortus putative outer membrane protein gene, partial
            cds; and putative 98 kDa outer membrane protein gene, complete cds.
ACCESSION   U72499
VERSION     U72499.1  GI:1657776 . . .

 to: cpn100396.seq  check: 452  from: 1  to: 3000

WPDEF
cpn100396
```

```
Symbol comparison table: /big1/gcg/gcgcore/data/rundata/nwsgapdna.cmp
CompCheck: 8760

         Gap Weight:      50      Average Match: 10.000
      Length Weight:       3      Average Mismatch:  0.000

            Quality:  14521            Length:    5027
              Ratio:  4.840             Gaps:      13
 Percent Similarity: 53.087   Percent Identity: 53.087

         Match display thresholds for the alignment(s):
                   | = IDENTITY
                   : =   5
                   . =   1

cpu72499.seq x cpn100396.seq March 2, 2001 11:29  ..
```

```
   1951 CAAGCAGTCCTCTCTAGAACGCAGGGGCCTCTTACCAAAAATGACAAAAA 2000
                                                         |
      1 ..............................................cgctct 6

   2001 CATGTCAAGTCTTTTCTGTATTAGACTTGCTTTCTTTTGTATCTTTGAGT 2050
          | |       || ||       |||| || |||    | || | ||
      7 tacctagtagaggttgagtgaatttcttgacttgtttctcctattggtgt 56

   2051 ATTTAACGAAACTGTTTTTTTCATTTCACCACA.AGTGTATAAAATGAGG 2099
          || |    |||   |   || |  |||    | |   | || ||||| |
     57 atctcttaaaatattaaattcaaaatcaaagtatatattttacaatgaag 106

   2100 CCTTCTTTATATAAGATTTTAATATCGTCGACGCTGACGTTACCAATATC 2149
          |||||||    ||| || || | ||   |      | || || | ||
    107 tcttctttccccaagtttgtattttctacatttgctattttccctttgtc 156
```

1

```
2150 TTTTCACTTCTCGCAATTGCATGCAGAAGTGGCTTTAACTCAAGAATCTA 2199
     | |         |         | |||   ||| || || |        |
 157 tatgattgctac......cgagacagttttggattcaagtgcgagtttcg 200


2200 TTCTCGATGCAAATGGAGCATTCAGTCCGCAATCTACAAGCACTGCGGGA 2249
     |     ||  ||||||   ||        |       |      || |
 201 atgggaataaaaatggtaatttttcagttcgtgagagtcaggaagatgct 250


2250 GGAACGATTTACAACGTCGAGAGTGATATTTCTATTGTAGATGTAGGACA 2299
     ||||| |  |||   |  || |  || |  || | || | | | || |
 251 ggaactacctacctatttaagggaaatgtcactctagaaatattcctgg 300


2300 GACAG...CGGCTCTTGCTTCCTCAGCTTTTGTTCAGACTGCAGACAACC 2346
     ||||   | ||  |   |        ||||   | |||    | |
 301 aacaggcacagcaatcacaaaaagctgtttttaacaacactaagggcgatt 350


2347 TAACTTTCAAAGGGAACAACCATAGCTTATCCATAACGAACGCGAATGCC 2396
     | ||||||||  |||  |||      |      |||    |   |  ||||
 351 tgactttcacaggtaacgggaactctctattgttccaaacggtggatgca 400


2397 GGAGCTAATCC...TGCGGGAATTAACGTTAACACTGCCGATAAGATTCT 2443
     ||  ||    |    || |  |||||  | |   |   ||||||     |
 401 gggactgtagcaggggctgctgttaacagcagcgtggtagataaatctac 450


2444 TACGCTGACAGATTTTTCTAAGTTGAGCTTTAAGGAATGCCCATCTTCTC 2493
     ||| | | || ||||||  | |     ||||   |   |  ||         |
 451 cacgtttatagggttttcttcgctatctttattgcgtctcctggaagtt 500


2494 TAGTGAATACTGGAAAAGGGGC...TATGAAATCCGGAGGAGCATTAAAC 2540
     | | ||| ||  ||||| ||  ||  ||      ||    || |
 501 cgataactaccggcaaaggagccgttagctgctctacgggtagcttgagt 550


2541 TTAGCGAATAATGCCAGTATTCTGTTTGATCAGAACTATTCCGCTGAGAA 2590
     ||   | ||  |||| ||||| |  | ||     | |||| ||| | || ||
 551 ttgacaaaaaatgtcagtttgctcttcagcaaaaacttttcaacggataa 600


2591 TGGTGGAGCCATCTCTTGCAAAGCTTTTTCTCTAACCGGCTCGAGCAAAG 2640
     ||| || || ||| |   ||| || |||| |||| ||  | | |
 601 tggcggtgctatcaccgcaaaaactctttcattaacagggactacaatgt 650


2641 AAATCAGCTTCACCACTAACTCTACTGCGAAAAAAGGTGGAGCGATTGCT 2690
     |       ||  |   || | | |  | || ||||| ||||| |||
 651 cagctctgttttctgaaaatacctcctcaaagaaaggcggagccattcag 700


2691 GCTACGGGAATAGCTCATCTTTCGGACAACCAAGGCACAATCAGATTTTC 2740
     ||  |         |    || | | ||||||||| | ||    |||||
 701 acttccgatgcccttaccattactggaaaccaaggggaagtctctttttc 750


2741 TGGGAACACTGCTGTGAATTCTGGGGGAGCAGTATATTCAGAAGCTTCTA 2790
     ||  || ||| ||  | |||||||| |  |||   |  | |||||||| ||
 751 tgacaatacttcttcggattctggagctgcaattttacagaagcctcgg 800


2791 TGACGATTGCAGGTAACAACCACGTTGCTTTTAGCAACAATGCTGTTTCC 2840
     |||| ||| |   |||     | |||| | |||| |||||  || |
 801 tgactatttctaataatgctaaagtttcctttattgacaataaggtcaca 850


2841 GG..............TTCATCTGATGGTTGCGGTGGAGCTATCCATTG 2875
```

2

```
       | |                    |      | | |    |  | |  | |  | | | | | | |     |
 851 ggagcgagctcctcaacaacggggggatatgtcaggaggtgctatctgtgc 900

2876 TAGCAAAACAGGTTCAGCACCGACCCTTACTATAAGAGATAACAAAGTCT 2925
       |    | | | | | |  | |  | | |    | |   | | |  | |  | | |   | |   | | |
 901 ttataaaactagtacagatactaaggtcaccctcactggaaatcagatgt 950

2926 TGATTTTTGAGGAAAATACTTCTTCAGCAAAAGGTGGAGCGATTTACACC 2975
       |  |  | |      | | | | | |  | |   | |  | |     | |  | | | | |  | |  | |
 951 tactcttcagcaacaatacatcgacaacagcgggaggagctatctatgtg 1000

2976 GATAAACTCATATTGACTTCTGGTGGGCCTACGGCATTTATCAATAACAA 3025
       |  | |  | | |    | | |  | | | |  | |  | |  |  | | |     | | |  |    |   | | |
1001 aaaaagctcgaactggcttccggaggacttaccctattcagtagaaatag 1050

3026 AGTTACCCATGCTACA...CCTAAGGGTGGAGCTATTGGTATTGCTGCCA 3072
       | |  |       |  | | | | |    | | | | | | | | | | | | | | |  | |  |  | | |  |    |    |
1051 tgtcaatggaggtacagctcctaaaggtggagccatagctatcgaagata 1100

3073 ATGGAGAATGTAGCTTAACCGCTGAACATGGGGATATTACTTTTGATAAT 3122
       | | |  | | | |   | |  | | |   | | | |  | |    | | |  | |  | | |      | | |
1101 gtggggaattgagtttatccgccgatagtggtgacattgtctttttaggg 1150

3123 AACCTGATGGCCACACAAGACAATGCTACAATAAAAAGAAATGCCATTAA 3172
       | |     |   |  | | |          | | |  | |    | | | | | | |     | |   |
1151 aa...tacagtcacttctactactcctgggacgaatagaagtagtatcga 1197

3173 CATTGAAGGCAATGGTAAATTCGTCAACTTACGTGCAGCGTCTGGAAAGA 3222
       |  |  |  |     |  | |  | |  |     |        | |  | | |  |  | |    | | | |  |
1198 cttaggaacgagtgcaaagatgacagctttgcgttctgctgctggtagag 1247

3223 CGATTTCTTTCTATGATCCTATCACAGTTG...........AAGGTAAT 3260
       |  | |  | |  | | | | | | | | | | | |  | | | |       |         | |  | |
1248 ccatctacttctatgatcccataactacaggatcatccacaacagttaca 1297

3261 GCTGCTGATCTTCTCACTTTGAATAAAGCTGAGGGTGATAAAACGTATAA 3310
       |  | |              |  |  |    | |  |     | |  | |     | |      |    | | | |
1298 gatgtcttaaaagttaatgagactccggcagattctgcactacaatatac 1347

3311 TGGAAGAATTATTTTTTCAGGAGAAAAGCTCACTGAAGAACAAGCTGCTG 3360
       | |  |    | |  | |    | |  | | | | | | | | | | |   | | |    | | |    | | | |  |
1348 agggaacatcatcttcacaggagaaaagttatcagagacagaggccgcag 1397

3361 TTGCGGATAACCTAAAGACAACATTTACACAGCCTATCACTTTAGCTGCT 3410
       |  |   | |  | | | | |    | |   |     | | | | | | | | |  | | | |   |   | |
1398 attctaaaaatcttacttcgaagctactacagcctgtaactctttcagga 1447

3411 GGTGAACTTGTGTTACGCAGCGGTGTGGAAGTAGAAGCAAAAACAGTCGT 3460
       | | |     | |      | | |         | |  | | |     |   |   |   |   |    | |  | |
1448 ggtactctatctttaaaacatggagtgactctgcagactcaggcattcac 1497

3461 GCAAACAGCAGGATCTTTGATTCTGATGGATGCAGGCACAAAGTTATCCG 3510
       | | |      | | | |   | | |     |       | | | | | | |  | | | | |    | |      | |    |
1498 tcaacaggcagattctcgtctcgaaatggacgtaggaactactcta...g 1544

3511 CAAAAACAGAAGATGCTACACTGACGAATCTGGCTATTAATCCGAATACC 3560
       |     |    | | |   |      | |  | |  | | | | | | | |    | | | | |      | | |
```

3

```
1545 aacctgctgatactagcaccataaacaatttggtcattaacatcagttct 1594
                                                            .
3561 TTAGATGGGAAAAAATTCGCCGTAGTCGATGCCGTTGCTGCTGGGAAGAA 3610
      |||| ||    |||      ||    | |·|| ||    ||| |    || ||
1595 atagacggtgcaaagaaggcaaaaatagaaaccaaagctacgtcaaaaaa 1644

3611 TGTGACTTTATCAGGTGCTATTGGCGTTATTGATCCTACAGGGAAGTTTT 3660
      | ||||||||||| ||  | ||    |  | || || || || | |||||
1645 tctgactttatctggaaccatcactttattggacccgacgggcacgtttt 1694

3661 ATGAAAACCATAAGCTAAATGATACGTTAGCTTTAGGAGGAATTCAACTT 3710
      ||||||| ||||   |||   || |    | |     |    | |  | ||
1695 atgaaaatcatagtttaagaaatcctcagtcctacgacatcttagagctc 1744

3711 TCTGGGAAAGGTTCGGTGACAACAACCAACGTGCCTAGTCATGTTGTTGG 3760
      |        ||  | ||  ||||   |||    ||| ||    || | |
1745 aaagcttctggaactgtaacaagcaccgcagtgactccagatcctataat 1794
                                  '
3761 TGTTGCTGAAACCCACTATGGTTATCAAGGAAACTGGTCTGTCAGTTGGG 3810
      -|-||- -|| -|+|- ||-|| -|||||-|†|--   --     |'||||||
1795 gggtgagaaattccattacggctatcaggga............acttggg 1832

3811 TCAAAGATAATAACTCTGATCCTAAAACACAAACAGCAATCTTTACCTGG 3860
      |   |  |   |               |  '  || |||| ||| | ||||
1833 gcccaattgtttggggggacaggggcttctacgactgcaaccttcaactgg 1882

3861 AATAAAACAGGATATGTTCCAAATCCTGAACGTCGTGCTCCGCTAGTACT 3910
      | ||||||| || |||  ||||  ||||| || |||   |    | |||| |
1883 actaaaactggctatattcctaatcccgagcgtatcggctctttagtccc 1932

3911 CAATAGCCTTTGGGGATCCTTTATAGATTTACGTTCTATTCAAGATGTCT 3960
      |||||| | |||     | |||||||||| | | ||| | ||  || |
1933 taatagcttatggaatgcatttatagatattagctctctccattatctta 1982
                                                            .
3961 TGGAACGTAGTGTTGATAGTATTCTTGAGACACGTCGTGGTTTGTGGGTC 4010
      ||||   |         ||  |  | |    || |   |||| ||| |||
1983 tggagactgcaaacgaagggttgc...agggagaccgtgcttttggtgt 2029

4011 TCTGGAATTGGGAACTTCTTCCATAAAGATCGGAATGCTGAAAATCGCAA 4060
      ||||| |    ||||||||||||||||| ||| | |        |   |||
2030 gctggattatctaacttcttccataaggatagtacaaaaacacgacgcgg 2079
                                                            .
4061 ATTCCGTCATATCAGTTCGGGATATGTGTTAGGAGCCACAACAAATACCT 4110
      || || ||| | |||    || ||||| ||||||| |   | ||||| |
2080 gtttcgccatttgagtggcggttatgtcataggaggaaacctacatactt 2129

4111 CGAGAGAGGATTCTCTTAGTGTGGCTTTCTGTCAGTTATTTGCAAAAGAT 4160
       |||   |    ||||||||| || || ||||||| | |||| || ||||
2130 gttcagataagattcttagtgctgcattttgtcagctctttggaagagat 2179

4161 AAAGACTACCTTGTAAGCAAGAACGCCGCAAACGTCTATGCGGGTTCTGT 4210
      | ||||||||| |||||     |||||   |  | |||||| | || || |
2180 agagactactttgtagctaagaatcaaggtacagtctacggaggaactct 2229

4211 ATATTATCAGCATGTGAGCAAGTTTGATGATCTCACGCGGTTATTTAATG 4260
      |||| |||||          |||      || || ||   |    | ||
2230 ctattaccagca......caacgaaacctatatctctcttccttgcaaac 2273
```

4

```
4261 GGCCTAACACGTGTTGTTCAGGGTTTTCTAAAGAGATTCCTATTTTCTTG 4310
        |        |  |  ||      ||  |||  ||||||||||||  || ||||
2274 tacggccttgttcgttgtcttatgttcctacagagattcctgttctcttt 2323

4311 GATGCACAAATTACCTATTGCCACACGGCCAACAACATGACAACGTCCTA 4360
        |  |   |||  |||      |||  ||||      |||  |  |||  |||         ||
2324 tcaggaaaccttagctacacccatacggataacgatctgaaaaccaagta 2373

4361 TACAGACTATCCTGAAGTGAAAGGTTCTTGGGGTAATGATACCCTGGGCT 4410
        ||||      ||||||      ||  ||||||           ||||||  |||||||      |  |       |
2374 tacaacatatcctactgttaaaggaagctgggggaatgatagtttcgctt 2423

4411 TAACTTTGTCTACTAGCGTACCTATCCCGGTATTTAGTTCTTCTATCTTT 4460
        ||     ||     |       ||  |    || ||·        !|   |                |  ||  |  |||
2424 tagaattcggtggaagagctccgatttgcttagatgaaagtgctctattt 2473

4461 GATAGTTATGCACCGTTTGCAAAATTACAAGTTGTCTATGCGCACCAAGA 4510
        ||          ||          ||  ||              |||||   ||    ||||||||||||   ||  ||  ||
2474 gagcagtacatgcccttcatgaaattgcagtttgtctatgcacatcagga 2523

4511 TGACTTTAAAGAACCAACAACAGAAGGCCGGGTCTTTGAAAGCAGCGATC 4560
         |    !||||||||||            ||||||||||   ||  |       ||||   |||   |||      ||
2524 aggttttaaagaacagggaacagaagctcgtgaatttggaagtagccgtc 2573

4561 TTCTCAACGTTTCTGTACCTATAGGTATAAAATTTGAGAAACTCTCCTAT 4610
        ||  |  ||    ||  |    |||||||  || ||       ||||||  ||           ||  |
2574 ttgtgaatcttgccttacctatcgggatccgatttgataaggaatcagac 2623

4611 GGAGAGAGAAGTGCTTATGATCTTACACTGATGTATATACCTGATGTGTA 4660
         |   |               |  ||     ||||  ||  ||            ||||                |||  |
2624 tgccaagatgcaacgtacaatctaactcttggttatactgtggatcttgt 2673

4661 CCGTCATAATCCAAGCTGTATGACAGGATTGGCGATCAATGACGTTTCCT 4710
         |||     |||  ||        ||||||  ||||    |  ||       ||  |    |       |  |||   |
2674 tcgtagtaaccccgactgtacgacaacactgcgaattagcggtgattctt 2723

4711 GGTTAACCACAGCTACGAATCTTGCTAGACAAGCTTTCATAGTTCGCGCG 4760
        ||   ||||      |  ||||||||  |  ||  |||||||||||      |      ||||   ||
2724 ggaaaaccttcggtacgaatttggcaagacaagctttagtccttcgtgca 2773

4761 GGTAACCATATTGCCTTAACCTCTGGTGTTGAGATGTTCAGTCAGTTTGG 4810
        ||  ||||||| ||    |||  |  |||         |  ||||          ||  ||   ||  |||
2774 gggaaccattttgctttaactcaattttgaagcctttagccaattttc 2823

4811 TTTCGAATTACGAAGCTCTTCAAGAAATTATAACGTAGATCTTGGCGCTA 4860
        |||  ||||||   ||     |  ||  ||     |  ||||||   ||  |||||  |   ||  ||  |
2824 ttttgaattgcgtgggtcatctcgcaattacaatgtagacttaggagcaa 2873

4861 AGGTCGCGTTCTAAAAAGCTCCCCCCTGCCCAGCTCTGGGCAGGGTTCTC 4910
         |    |      ||||||   |  |     |     |     ||   ||         |
2874 aataccaattctaatgcgttagctttggtaaagagctccatacatcgaag 2923

4911 CTTATTCTAGTCTAGA................................ 4926
         |         |  |   |
2924 ggaaaagagcttttaagatttcttgaaggctcttttcgatttcgatttcc 2973
```

5

Alignment 2

```
GAP of: cpu65942.seq   check: 1642   from: 1  to: 6110

WPDEF        Chlamydophila abortus strain S26/3 POMP91A and POMP90A precursor,
LOCUS        CPU65942       6110 bp      DNA                BCT        07-MAY-1999
DEFINITION   Chlamydophila abortus strain S26/3 POMP91A and POMP90A precursor,
             genes, complete cds.
ACCESSION    U65942
VERSION      U65942.1  GI:1617505 . . .

to: cpn100396.seq   check: 452   from: 1  to: 3000

WPDEF
cpn100396


Symbol comparison table: /big1/gcg/gcgcore/data/rundata/nwsgapdna.cmp
CompCheck: 8760

        Gap Weight:      50    Average Match: 10.000
     Length Weight:       3    Average Mismatch:  0.000

           Quality:  13146          Length:   6140
             Ratio:  4.382            Gaps:     20
Percent Similarity: 49.428    Percent Identity: 49.428


        Match display thresholds for the alignment(s):
                  | = IDENTITY
                  : =    5
                  . =    1

cpu65942.seq x cpn100396.seq March 2, 2001 13:39  ..
```

```
2251 CTTTAAGGAAAACAATAGCGATCAAGGAAGATACTTTGAAAGTAATAACC 2300
                   || ||   |  ||  |||| | ||
   1 ...............cgctcttacctagtagaggttgagtgaatttctt 34

2301 TCACCAACTTATCTATGCCCATAGGTGTGAAATTGGAAAAATTCTCTCAC 2350
            |   ||||    ||  | | ||| |   ||| |
  35 gacttgtttctcctattggtgtatctcttaaaatattaaattcaaaatca 84

2351 AAGGATACCGCTTCTTACAACCTTACCCTGGCTTATGCTCCAGATATCGT 2400
     ||| |||    ||   | | |   || | | |||     | ||
  85 aagtatatattttacaatgaagtcttctttccccaagtttgtattttcta 134

2401 AAGAAGCAACCCTGACTGTACTGCCTCTCTGTTAGTGAGCCCAACCTCTG 2450
       |   ||       ||| | || ||| ||| | |||
 135 catttgc..........tattttccctttgtctatgattgctaccgaga 173

2451 CTGTCTGGGTAACTAAAGCTAATAACCTTGCGCGGCATGCTTTCATATTA 2500
      | || | ||   | |  || | | | ||| |   | |   | ||
 174 cagttttggattcaagtgcgagtttcgatgggaataaaaatggtaatttt 223
```

```
2501 CAAGCAGGAAATTACTTGGCTTTAACTCGCAACACGGAGCTCTTCAGTCA 2550
          ||   | |      ||      ||| |  || | ||    ||| |
 224 tcagttcgtgagagtcaggaagatgctggaactacctacct....attta 269

2551 ATTCGGTTTCGAACTCCGAGGTTCTTGCCGCACCTATAACATAGATCTCG 2600
      |     |  |     ||   | || | | |     |||    ||| |
 270 agggaaatgtcactctagaaatattcctggaacaggcacagcaatcaca 319

2601 GATCGAAGATCCAGTTCTAATCTCATCCACCTCCCCTGCTCCGTAGACAC 2650
      | |  |      |||| |   || |  |      | |    || |||
 320 aaaagctg.......ttttaacaacactaagggcgatttgactttcacag 362

2651 GGAGCAGGGCCTCCCTTTGTCTACTCTAGCTGGTGACCTTGACTTTTAAT 2700
      | | | ||  ||| ||  | |   |   | |||   |  ||||| |
 363 gtaacgggaactctctattgttccaaacggtggatgcagggactgtagca 412

2701 TATTTTTATGATCGAGTTAGGATATCTCAAATCACCCAAAGCATGAATTT 2750
       | ||| |            |               | || |  |
 413 ggggctgctgtt.........................aacagcagcgt 435

2751 ATGAAACATCCAGTCTACTGGTTCTTAATATCCTCGAGCCTATTTGCCTC 2800
       | | |       | ||    |    | | ||    | |||   |
 436 ggtagataaatctaccacgtttatagggttttcttcgctatctttattg 485

2801 GAATTCTTTGAGCTTCGCTAACGACGCTCAAACAGCCTTAACTCCCTCCG 2850
       || | ||  |||| |||| ||| ||   ||||   |    |    |
 486 cgtctcctggaagttcgataactaccggcaaaggagccgttagctgctct 535

2851 ATAGCTATAATGGAAATGTGACCTCTGAGGAGTTCCAGGTAAAAGAAACT 2900
      | | || ||    ||    |    | | ||           | | | ||
 536 acgggtagcttgagtttgaca............aaaaatgtcagtttgc 572

2901 TCATCAGGAACAACGTATACTTGTGAAGGCAATGTGTGTATCTCCTTTGC 2950
      || |||| || ||| |      ||      |  | ||||   |    ||
 573 tcttcagcaaaaactt....ttcaacggataatggcggtgctatcaccgc 618

2951 AGGGAAAGATTCAGGTCTAAAGAAAAGTTGTTTCTCAGCTACTGATAACC 3000
      |  |    ||||    |  ||  |   | |   |       |  | |
 619 aaaaactctttcattaacagggactacaatgtcagctctgttttctgaaa 668

3001 TTACCTTCCTAGGAAACGGGTATACTCTTTGCTTTGATAATATTACTACT 3050
      |||| |  |   | ||  ||     | || |   | ||    ||  |||
 669 atacctcctcaaagaaaggcggagccattcagacttccgatgcccttacc 718

3051 ACAGCTAGTAACCCCGGAGCCATTAATGTTCAAGGTCAAGGAAAAACCTT 3100
      |   || | |||| || || |   |  | || | |    |
 719 attactggaaaccaaggggaagtctcttttctgacaa............ 756

3101 AGGCATCTCAGGATTTTCTTTATTTTCATGTGCTTATTGTCCTCCAGGCA 3150
                ||||| ||| | |||    | | |||
 757 ............tacttcttcggattctggagctgcaatttttacagaag 794

3151 CAACTGGTTACGGAGCTATACAGACTAAAGGCAACACAACTTTAAAAGAT 3200
      | |          |||   | ||||||   |  || | |  ||  |
 795 cctcggtgactatttctaataatgctaaagtttcctttattgacaataag 844

3201 AACTCTAGTCTTGTCTTCCATAAAAACTGCTCAACAGCAGAAGGTGGGGC 3250
```

7

```
                |  |  |         ||  |    |  |||    |   · |    |||  |||||
    845 gtcacaggagcgagctcctcaacaacggggggatatgtcaggaggtgctat 894

   3251 TATCCAGTGTAAAGGAAGCAGTGATGCTGAATTAAAAATAGAAAATAATC 3300
             |  |||||    ||  |   |||  |·|  |   |  |    |        ||||
    895 ctgtgcttataaaactagtacagatactaaggtcaccctcactggaaatc 944

   3301 AGAATCTGGTTTTCTCAGAAAACTCCTCCACTTCAAAAGGCGGGGCTATT 3350
             |||   |   |  |||      |  ||   |  |  ||   ||    ||  ||  |||||
    945 agatgttactcttcagcaacaatacatcgacaacagcgggaggagctatc 994

   3351 TATGCTGATAAACTCACCATTGTCTCAGGTGGGCCTACATTATTTTCTAA 3400
             ||||     |  ||  |||      |  |   ||  ||  ||  |  |||  ||||      ||
    995 tatgtgaaaaagctcgaactggcttccggaggacttaccctattcagtag 1044

   3401 CAACTCTGTATCCAACGGTTCATCCCCTAAAGGCGGAGCTATTAGCATAA 3450
             ||    |||          |||  ||  |  |||||||||  |||||  ||       ||
   1045 aaatagtgtcaatggaggtacagctcctaaaggtggagccatagctatcg 1094

   3451 AAGATTCAAGTGGTGAATGTAGCCTAACCGCTGATCTCGGAGATATTACC 3500
             |||||      |||||  ||||  ||   ||   |||||  |||     ||  ||  |||     |
   1095 aagat...agtgggggaattgagtttatccgccgatagtggtgacattgtc 1141

   3501 TTCGATGGGAACAAAATCATCAAAACTAGTGGTGGAAGTTCTACAGTAAC 3550
             ||       |||||  |  |  |||      ||||  |   |||            ||      |
   1142 tttttagggaatacagtcacttctactactcctgg......gacgaatag 1185

   3551 AAGAAATTCCATAGATCTCGGCACAGGGAAATTTACAAAGCTACGTGCTA 3600
             |||  |  |  |         |   |   ||   |   |||     |  |||  ||
   1186 aagtagtatcgacttaggaacgagtgcaaagatgacagctttgcgttctg 1235

   3601 AAGACGGCTTCGGAATTTTCTTCTATGACCCTATTACTGGGGGAGGAT.. 3648
             |   ||     |   ||  |  |||||||||  ||  ||  |||      |||   ||
   1236 ctgctggtagagccatctacttctatgatcccataactacaggatcatcc 1285

   3649 .........CTGATGAACTAAACATTAATAA.........AAAAGAAAC 3679
                      |  |||||     ||||   |||||| |          |  |      |
   1286 acaacagttacagatgtcttaaaagttaatgagactccggcagattctgc 1335

   3680 TGTTGATTATACAGGAAAGATCGTCTTCTCAGGTGAAAAATTATCCGATG 3729
             |   |  |||||||||  ||  |||  |||||  ||   |||||  |||||  ||
   1336 actacaatatacagggaacatcatcttcacaggagaaaagttatcagaga 1385

   3730 AAGAAAAAGCACGAGCGGAAAACCTAGCTTCTACTTTCAACCAACCCATC 3779
             |||       |||      |   ||||  ||   ||||  |     |        ||  ||   |
   1386 cagaggccgcagattctaaaaatcttacttcgaagctactacagcctgta 1435

   3780 ACATTATCAGCAGGATCTCTTGTACTTAAAGATGGTGTATCTGTAACCGC 3829
             ||   |  |||||  |||   ||||       |  |||  ||||| ||  ||  |              |
   1436 actctttcaggaggtactctatctttaaaacatggagtgactctgcagac 1485

   3830 AAAACAAGTAACGCAGGAAGCGGGATCTACCGTTGTCATGGATCTAGGGA 3879
             |     |  |  ||  ||   ||||  |   |||      |  |   |||||   ||||  |
   1486 tcaggcattcactcaacaggcagattctcgtctcgaaatggacgtaggaa 1535

   3880 CCACATTACAGACGCCTTCTTCAGGTGGAGAAACCATCACCCTAACTAAT 3929
             |  ||   ||  |               |   ||   ||  ||  |  ||||| |||     |||
```

8

```
1536 ctactctaga...........acctgctgatactagcaccataaacaat 1573
                .              .              .          
3930 CTAGATATTAACATCGCCTCGTTGGGGGGGGGGGGGGGTACCTCTCCTGC 3979
     | |  |||||||||   || |          | ||| |      ||
1574 ttggtcattaacatcagttctat.........agacggtgcaaagaaggc 1614

3980 TAAACTCGCAACAAATACAGCAAGTCAAGCTATAACTAT...TAACGCTG 4026
     ||| | | ||| || | |      || | | ||| |   |     |
1615 aaaaatagaaaccaaagctacgtcaaaaaatctgactttatctggaacca 1664

4027 TCAATCTAGTCGATGCTGATGGCAATGCTTATGAAGATCCTATTCT.... 4072
     ||| | || | || |    |||| ||||||| ||| || | |
1665 tcactttattggacccgacgggcacgttttatgaaaatcatagtttaaga 1714

4073 ..TGCTACGTCTAAACCTTTCACAGCAATAGTAGCTACAACTAACGCTAG 4120
       | || ||| |    || || | |||| |   | | |
1715 aatcctcagtcctacgacatcttagagctcaaagcttctggaactgtaac 1764

4121 TACAGTCACACAGCCTACAGATAATCTAACAAATTATGTCCCTCCTACTC 4170
     | -   |=|| = | =| ||=|| | ||| - | |              |
1765 aagcaccgcagtgactccagatcctataatgggtga......gaaattcc 1808

4171 ATTACGGTTACCAAGGAAATTGGACAGTAACTTGGGACACCGAAACAGCT 4220
     ||||||| || || |||| ||||     || |        |   |||
1809 attacggctatcagggaacttggggcccaattgtttgggggacaggggct 1858

4221 ACAAAAACAGCCACTCTAACTTGGGAACAAACTGGCTACTCCCCTAACCC 4270
     | |  || || || | |  |||   |||||||||||   ||||| ||
1859 tctacgactgcaaccttcaactggactaaaactggctatattcctaatcc 1908

4271 AGAACGTCAAGGACCTTTAGTCCCGAATACTCTTTGGGGTGCATTCTCTG 4320
     || |||   ||  |||||||||| ||||   | |||  ||||||     |
1909 cgagcgtatcggctctttagtccctaatagcttatggaatgcatttatag 1958

4321 ACCTCAGAGCTATACAAAACTTAATGGATATTAGCGTCAATGGCGCTGAC 4370
     |  | ||  || | || |  | ||||||  | |    |  || |
1959 atattagctctctccattatcttatggagactgcaaacgaagggttgcag 2008

4371 TACCATAGAGGTTTTTGGGTATCCGGTCTAGCTAACTTCTTACACAAAAG 4420
      |  | | |||||||   | ||  || |||||||||||| || ||
2009 ggagaccgtgctttttggtgtgctggattatctaacttcttccataagga 2058

4421 TGGCTCTGATACTAAACGCAAGTTCCGTCACAATAGCGCCGGATACGCTT 4470
     | |  | | ||  ||||  ||| || ||     || | ||| || |
2059 tagtacaaaaacacgacgcgggtttcgccatttgagtggcggttatgtca 2108

4471 TAGGCGTCTACGCAAAAACTCCTTCTGATGATATTTTCAGTGCGGCTTTC 4520
     |||| |   ||  | | ||| ||| ||| | ||| | ||||| || ||
2109 taggaggaaacctacatacttgttcagataagattcttagtgctgcattt 2158

4521 TGCCAACTCTTCGGAAAGGACAAAGACTATTTAGTGTCGAAAAACAACGC 4570
     || || |||||| ||||  || | ||||||| || ||  | || || | |
2159 tgtcagctctttggaagagatagagactactttgtagctaagaatcaagg 2208

4571 CAACATTTACGCAGGTTCTCTCTATTATCAGCATATCTCCTATTGGAGCG 4620
     |    | |||| |||  |||||||||| |||||| | |      |  | |
2209 tacagtctacggaggaactctctattaccagcacaacgaaacctatatct 2258
```

9

```
4621 CTTGGCAGAATCTGCTACAAAACACTATCGGTGCAGAAGCTCCGTTAGTC 4670
     ||  |      ||||        |     |  |  | |||   ||
2259 ctcttccttgcaaactacggccttgttcgttgtcttatgttcctacagag 2308

4671 CTTAACGCACAGTTAACTTATTGTCATGCTTCAAACGACATGAAAACCAA 4720
     ||   |  |  ||  |      ||  |  | | | | | | | | | |
2309 attcctgttctcttttcaggaaaccttagctacacccatacggataacga 2358

4721 CATGACGACTACTTACGCTCCTCGTAAAACAACGTATGCAGAAATCAAGG 4770
     |||   || |  |            | |||||| ||| |     | || |
2359 tctgaaaaccaagt.............atacaacatatcctactgttaaag 2396

4771 GTGATTGGGGTAACGATTGTTTCGGAGTCGAGCTTGGTGCAACTGTGCCT 4820
     |     |||||  ||  ||| |||||||   | ||  | |||||  |  ||
2397 gaagctgggggaatgatagtttcgctttagaattcggtggaagagctccg 2446

4821 ATCCAAACAGAATCTTCTCTCCTATTCGATATGTACTCACCTTTCCTGAA 4870
     ||       |||      |     ||||||  ||    || ||| ||||
2447 atttgcttagatgaaagtgctctatttgagcagtacatgcccttcatgaa 2496

4871 GTTTCAACTTGTGCATACGCACCAAGATGACTTTAAGGAAAACAATAGCG 4920
     || ||  |||| ||  | || || || |  |||||| |||  |    | |
2497 attgcagtttgtctatgcacatcaggaaggtttttaaagaacagggaacag 2546

4921 ATCAGGGAAGATACTTCGAAAGCAGCAATCTCACCAACCTTTCTCTGCCT 4970
      ||  |   |  | | |  ||| |||   |||   || ||| |   | |||
2547 ...aagctcgtgaatttggaagtagccgtcttgtgaatcttgccttacct 2593

4971 ATCGGCATCAAGTTTGAGAGATTTGCTAACAACGATACAGCTTCTTATCA 5020
     ||||| |||   ||||| |     | || | |     || | ||  |
2594 atcgggatccgatttgataaggaatcagactgccaagatgcaacgtacaa 2643

5021 TGTCACTGCTGCTTATTCTCCTGATATCGTAAGAAGTAACCCTGACTGTA 5070
     | | ||| || |||| ||   ||| | || | ||||||||| |||||||
2644 tctaactcttggttatactgtggatcttgttcgtagtaacccccgactgta 2693

5071 CTACTTCTCTGTTAGTAAGCCCCGACTCTGCTGTCTGGGTAACGAAAGCC 5120
     | ||  | ||| | | |||    || |||        ||| |||    |
2694 cgacaacactgcgaattagcggtgattct......tggaaaaccttcggt 2737

5121 AACAACCTTGCGCGAAGCGCCTTCATGCTACAAGCAGGAAACTACTTGTC 5170
     |  ||  | ||  ||     || ||  | || | || ||||| ||| | ||  |
2738 acgaatttggcaagacaagctttagtccttcgtgcagggaaccattttttg 2787

5171 TTTAAGTCACAACATAGAAATCTTCAGCCAGTTCGGTTTCGAGCTCAGGG 5220
     || |     || | |||  ||| |||||| ||    ||| || |  ||
2788 ctttaactcaaattttgaagcctttagccaattttcttttgaattgcgtg 2837

5221 GATCTTCACGAACCTATAACGTAGATCTCGGATCGAAGATCCAGTTCTAA 5270
     | || || ||   ||  || |||||| | ||| | ||   ||| ||||||
2838 ggtcatctcgcaattacaatgtagacttaggagcaaaataccaattctaa 2887

5271 TCTCATCCACCTCTCCTACCCCGTTCCCACGGGGTAGGGCCTTATAACCT 5320
     | | | ||  | |    ||   ||| | ||||   ||||          |
2888 t.gcgttagctttggtaaagagctccatacatcgaagggaaaagagcttt 2936
```

10

```
5321 TGACTCTCCTTGGAATATCCCGAACAAAGAGTCAAAAAAATACGTCTCAT 5370
     | |    | |||| |   ||      |      |      |  |  || |   |
2937 taagatttcttgaaggctcttttcgatttcgatttccattttagtgtttt 2986

5371 AGGATTCTGGTCTATTCCCTTTTCTTTATTTTGGAGAAAAATCTGATCAC 5420
        |      | |
2987 gctaaaacactttc.................................... 3000
```

Alignment 3

```
GAP of: pomp90.pep   check: 1153  from: 1  to: 839

WPDEF      prt
pomp90 prt

 to: 396prt.pep  check: 2118  from: 1  to: 928

WPDEF
396prt

 Symbol comparison table: /big1/gcg/gcgcore/data/rundata/blosum62.cmp
 CompCheck: 6430
BLOSUM62 amino acid substitution matrix.
Reference: Henikoff, S. and Henikoff, J. G. (1992). Amino acid
           substitution matrices from protein blocks.  Proc. Natl. Acad.
           Sci. USA 89: 10915-10919.


        Gap Weight:        8      Average Match:   2.912
     Length Weight:        2   Average Mismatch:  -2.003

          Quality:    1485          Length:     958
            Ratio:   1.770            Gaps:      29
Percent Similarity: 54.512   Percent Identity: 47.960


       Match display thresholds for the alignment(s):
                 | = IDENTITY
                 : =    2
                 . =    1


pomp90.pep x 396prt.pep    March 2, 2001 16:59  ..



     1 MKHPVYWFLISSSLFASNSLSFANDAQTALTPSDSYNGNVTSEEFQVKET 50
       ||     |.|.  ||   ||      :| |  | |:.||  .  | |:|.
     1 mkssfpkfvfst..faifplsmi.atetvldssasfdgn.kngnfsvres 46

    51 ..SSGTTYTCEGNVCI.SFAGKDSGLKKSCFSAT.DNLTFLGNGYTLCFD 96
        .||||  .||| : .   |   . : ||||. |  .||| ||| .| |
    47 qedagttylfkgnvtlenipgtgtaitkscfnntkgdltftgngnsllfq 96

    97 NITTTASNPGAINVQGQGKTLGISGFSLFSCAYCPPG..TTGYGAIQ.TK 143
       :          |:|        |.    ||| |      ||| ||:  .
    97 tvdagtvagaavnssvvdksttfigfsslsfiaspgssittgkgavscst 146

   144 GNTTLKDNSSLVFHKNCST.........................AEG 165
       |. .|  | ||.| || ||                          .|
   147 gslsltknvsllfsknfstdnggaitaktlsltgttmsalfsentsskkg 196

   166 GAIQ......................................... 169
       ||||
   197 gaiqtsdaltitgnqgevsfsdntssdsgaaifteasvtisnnakvsfid 246

   170 ...................C..KGSSDAELKIENNQNLVFSENSSTSKGG 198
                          |  | |.|  .. :  || |.|| |.||. ||
   247 nkvtgasssttgdmsggaicayktstdtkvtltgnqmllfsnntsttagg 296
```

12

```
199 AIYADKLTIVSGGPTLFSNNSVSNGSSPKGGAISIKDSSGECSLTADLGD 248
    |||  || : |||  |||| |||.  |..||||||.|.| |||  ||.|| ||
297 aiyvkklelasggltlfsrnsvnggtapkggaiaied.sgelslsadsgd 345

249 ITFDGNKIIKTSGGSSTVTRNSIDLGT.GKFTKLRAKDGFGIFFYDPITG 297
    | | || :  |. |..   |.||||||  | | ||.   |  |:||||||
346 ivflgntvtsttpgtn...rssidlgtsakmtalrsaagraiyfydpitt 392

298 GGS....DELNINK...KETVDYTGKIVFSGEKLSDEEKARAENLASTFN 340
    | |    | | :|.      . ||| |:|.|||||: | |  ..|| |
393 gssttvtdvlkvnetpadsalqytgniiftgeklseteaadsknltskll 442

341 QPITLSAGSLVLKDGVSVTAKQVTQEAGSTVVMDLGTTLQTPSSGGETIT 390
    ||:||| |.| || ||..    ||:| | . ||.|||: |.     :| |
443 qpvtlsggtlslkhgvtlqtqaftqqadsrlemdvgttle.pa...dtst 488

391 LTNLDINIASLGGGGGTSPAKLATNTASQAITIN.AVNLVDADGNAYEDP 439
    : || |||.|:    |   ||: |   |. :|:.  : |.| |  ||.
489 innlvinissi...dgakkakietkatsknltlsgtitlldptgtfyenh 535

440 ILATSKPFTAI...VATTNASTVTQPTDNLTNYVPPTHYGYQGNW.TVTW 485
    |    . :  :    . | ||   | | :    ||||||  |  :  |
536 slrnpqsydilelkasgtvtstavtp.dpimg..ekfhygyqgtwgpivw 582

486 DTETATKTATLTWEQTGYSPNPERQGPLVPNTLWGAFSDLRAIQNLMDIS 535
    |   .| |||  |  .|||  |||||  |  ||||.|| ||  |:  .:  ||:  .
583 gtgast.tatfnwtktgyipnperigslvpnslwnafidisslhylmeta 631

536 VNGADYHRGFWVSGLANFLHKSGSDTKRKFRHNSAGYALGVYAKTPSDDI 585
    |      | ||  .||.|| ||   .  |:|  |||  |  || :|    | ||| |
632 neglqgdrafwcaglsnffhkdstktrrgfrhlsggyviggnlhtcsdki 681

586 FSAAFCQLFGKDKDYLVSKNNANIYAGSLYYQH....ISYWSAWQNLLQN 631
    |||||||||:|:|| |.||   :| |.||||||      ||
682 lsaafcqlfgrdrdyfvaknqgtvyggtlyyqhnetyislpcklrpcsls 731

632 TIGAEAPLVLNAQLTYCHASNDMKTNMTTTYAPRKTTYAEIKGDWGNDCF 681
    :   | |.. .  |.| |  ||:||      : |||  :|| |||| |
732 yvpteipvlfsgnlsythtdndlkt.:.......kyttyptvkgswgndsf 773

682 GVELGATVPIQTESSLLFDMYSPFLKFQLVHTHQDDFKENNSDQGRYFES 731
    .| |  || : | ||: | ||:| | |: ||: |||  . : | |
774 alefggrapicldesalfeqympfmklqfvyahqegfkeqgt.earefgs 822

732 SNLTNLSLPIGIKFERFANNDTASYHVTAAYSPDIVRSNPDCTTSLLVSP 781
    | | ||.|||||:|:: ..    |.|..|  |. |:||||||||||.| :|
823 srlvnlalpigirfdkesdcqdatynltlgytvdlvrsnpdctttlrisg 872

782 DSAVWVTKANNLARSAFMLQAGNYLSLSHNIEIFSQFGFELRGSSRTYNV 831
    ||  ||    |||| | .|.|||:   . | | |||| |||||||| |||
873 ds..wktfgtnlarqalvlragnhfcfnsnfeafsqfsfelrgssrnynv 920

832 DLGSKIQF 839
    |||.| ||
921 dlgakyqf 928
```

13

Alignment 4

```
GAP of: pomp91.pep   check: 2232   from: 1  to: 847

WPDEF          prt
pomp91 prt

 to: 396prt.pep   check: 2118   from: 1  to: 928

WPDEF
396prt


 Symbol comparison table: /big1/gcg/gcgcore/data/rundata/blosum62.cmp
 CompCheck: 6430
BLOSUM62 amino acid substitution matrix.
Reference: Henikoff, S. and Henikoff, J. G. (1992). Amino acid
           substitution matrices from protein blocks.  Proc. Natl. Acad.
           Sci. USA 89: 10915-10919.

           Gap Weight:      8      Average Match:   2.912
        Length Weight:      2      Average Mismatch: -2.003

              Quality:   1468             Length:    967
                Ratio:   1.733              Gaps:     29
   Percent Similarity: 51.238    Percent Identity: 44.059

        Match display thresholds for the alignment(s):
                    | = IDENTITY
                    : =    2
                    . =    1

 pomp91.pep x 396prt.pep   March 2, 2001 17:01  ..
```

```
   1 ...............................MKHPVYWFLISS 12
                                    .  . . :  .
  51 gttylfkgnvtlenipgtgtaitkscfnntkgdltftgngnsllfqtvda 100

  13 SLIVSNSLYSEEPDQKTLTSAHSYNGNTNSEPFNPLSTSNSNGTIYTCTG 62
      :    .. |   |. |      | . .  . | .:. |    | :   ||
 101 gtvagaavnssvvdksttfigfs.slsfiaspgssitt..gkgavscstg 147

  63 NICIAYAGLDGSGLSSSCF.TD.....TAGNLSFLGNGYTLCFDNITTQS 106
     .: :      . | | | | ||    || || |  . |   | |
 148 slslt...knvsllfsknfstdnggaitaktlsltgttmsalfse.ntss 193

 107 SHPGAISVSGTNKTLDISGFSLFSCAYCPPGATGYG.AIKAVGNTTIKDN 155
        |||     |. | |.| .    .:   .. | ||   . || .|
 194 kkggaiq...tsdaltitg.nqgevsfsdntssdsgaaifteasvtisnn 239

 156 SSLVFHKNCSTG.........EGGAIQCKASSSEAELKIENNQNLVFAEN 196
       . . |  |  ||         ||||    .|.: .. :  || |.|. |
 240 akvsfidnkvtgasssttgdmsggaicayktstdtkvtltgnqmllfsnn 289
```

14

```
197 SSSSSGGAIYADKLTIVSGGPTLFSNNSVS.ASSPKGGAICIKDSGGECS 245
    ·|...|||||| ||  : ||| |||| |||·  ··||||||| |·|| || |
290 tsttaggaiyvkklelasggltlfsrnsvnggtapkggaiaieds.gels 338

246 LTADLGDITFDGNKIIKTNGGSPTVTRNSIDLGSSGKFTKLNAKEGFGIF 295
    |·|| ||| | ||  : |   ·|   |·|||||·| | | |  ·  | |:
339 lsadsgdivflgntvtst...tpgtnrssidlgtsakmtalrsaagraiy 385

296 FYDPITGGGS....DELNINK...QDTVDYTGKIVFSGERLSDEEKKVAA 338
    ||||||| ||    | | :|·   · ||| |:|·||:||: |   ·
386 fydpittgssttvtdvlkvnetpadsalqytgniiftgeklseteaadsk 435

339 NLKSDFKQPLKIGSGSLILKDGVTLETKSFTQTEGATVVMDLGTTLQTPS 388
    || |   ||·  : |·| || ||||:|··|||   · · ||·||||: |·
436 nltskllqpvtlsggtlslkhgvtlqtqaftqqadsrlemdvgttle.pa 484

389 SGGETTTLTNLDINVASLGGGGVAPDPAKVEATTESKTVTIN.AVNLVDD 437
    :| |: || ||:·|: |        ||:|    || ·|:·  : |·|
485 ...dtstinnlvinissidga....kkakietkatsknltlsgtitlldp 527

438 NGNAYEYPILAASQPFTAIEVRSGSSGSITKPTTNLENYT.PPTHYGYQG 486
    |  ||   |   | : :|·: ·||·:|        :      ||||||
528 tgtfyenhslrnpqsydilelk..asgtvtstavtpdpimgekfhygyqg 575

487 NW.TVTWKQGSSAQEKTATLTWEQTGYSPNPERQGSLVPNTLWGSFSDIR 535
    |  : |  |·|     ||| | ·||| ||||| |||||||·|| ·| ||
576 twgpivwgtgast...tatfnwtktgyipnperigslvpnslwnafidis 622

536 AIQNLMDISVNGADYHRGFWVSGLGNFLHKSGSDTKRKFRHNSAGYALGV 585
    ·:  ||: ·  |    | || ·|| || ||   · |:| ||| | || :|
623 slhylmetaneglqgdrafwcaglsnffhkdstktrrgfrhlsggyvigg 672

586 YAQTPSEDVFSAAFCQLFGKDKDYLVSKNSSTVYAGSIYYQHISYWNTWN 635
    | |:  : ||||||||:|:|| |·|| ||| |·:||||    :  ·
673 nlhtcsdkilsaafcqlfgrdrdyfvaknqgtvyggtlyyqhnetyislp 722

636 TLLQ....NTLGAEAPLVLNAQLAYCHASNNMKTNMTDTYAPPKTTYSEI 681
    |·     · · | |··· |·| | |·:||    |    |||  :
723 cklrpcslsyvpteipvlfsgnlsythtdndlktk....y....ttyptv 764

682 KGDWGNDCFGVEFGAKAPI.ETASLLFDMYSPFVKLQLVHAHQDDFKENN 730
    || |||| |·||| :|||    | ||: | ||·|| |:|||: |||
765 kgswgndsfalefggrapicldesalfeqympfmklqfvyahqegfkeqg 814

731 SDQGRYFESNNLTNLSMPIGVKLEKFSHKDTASYNLTLAYAPDIVRSNPD 780
    · : ||| |· | ||·:|||:: :| |    |·||||| |  |:VRSNPD
815 t.earefgssrlvnlalpigirfdkesdcqdatynltlgytvdlvrsnpd 863

781 CTASLLVSPTSAVWVTKANNLARHAFILQAGNYLALTRNTELFSQGFGEL 830
    || ·| :| | ||   |||| | :|·|||:    | | |||| |||
864 ctttlrisgds..wktfgtnlarqalvlragnhfcfnsnfeafsqfsfel 911

831 RGSCRTYNIDLGSKIQF 847
    ||| | ||:|||·| ||
912 rgssrnynvdlgakyqf 928
```